Fully-fledged macromolecular crystallography under high pressure and lessons: plea for ultra-short wavelength beamlines.

R. Kahn	IBS, Grenoble, France
R. Fourme & E. Girard	Synchrotron-SOLEIL, Saint Aubin, France
I. Ascone	LURE, Orsay, France
M. Mezouar	ESRF, Grenoble, France
T. Lin & J.E. Johnson	The Scripps Research Institute, La Jolla, CA, USA

1 bar ≈ 0.1 MPa

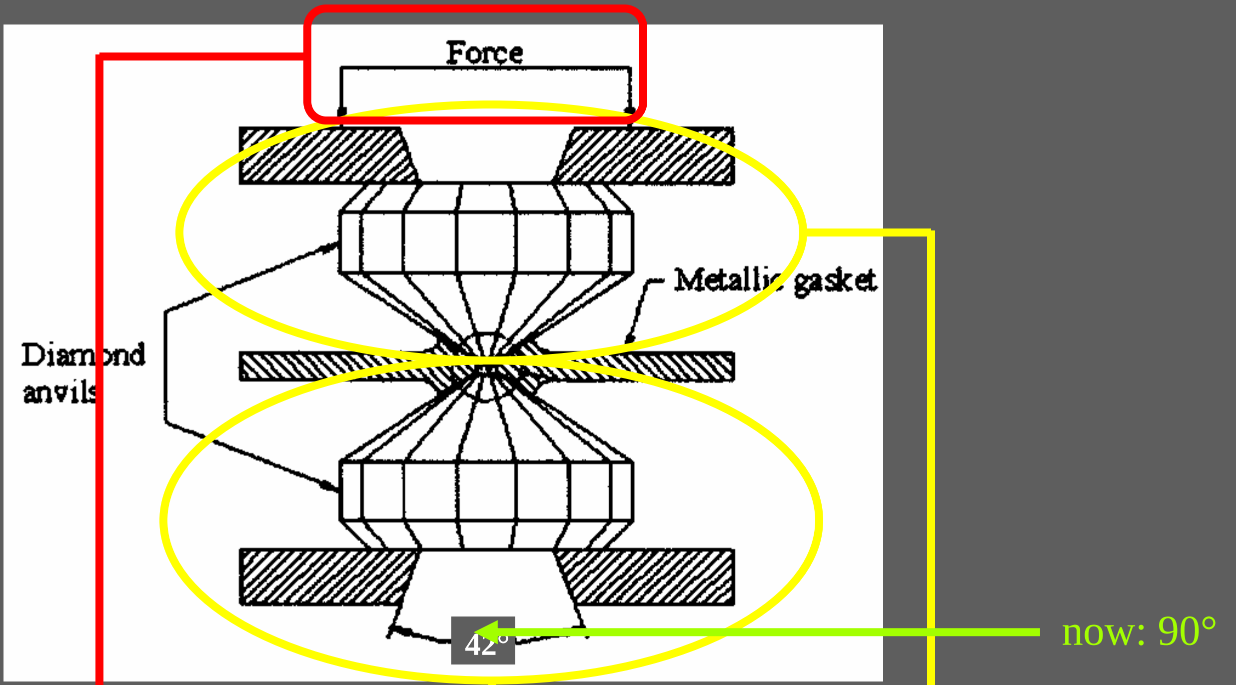
1 kbar ≈ 100 MPa

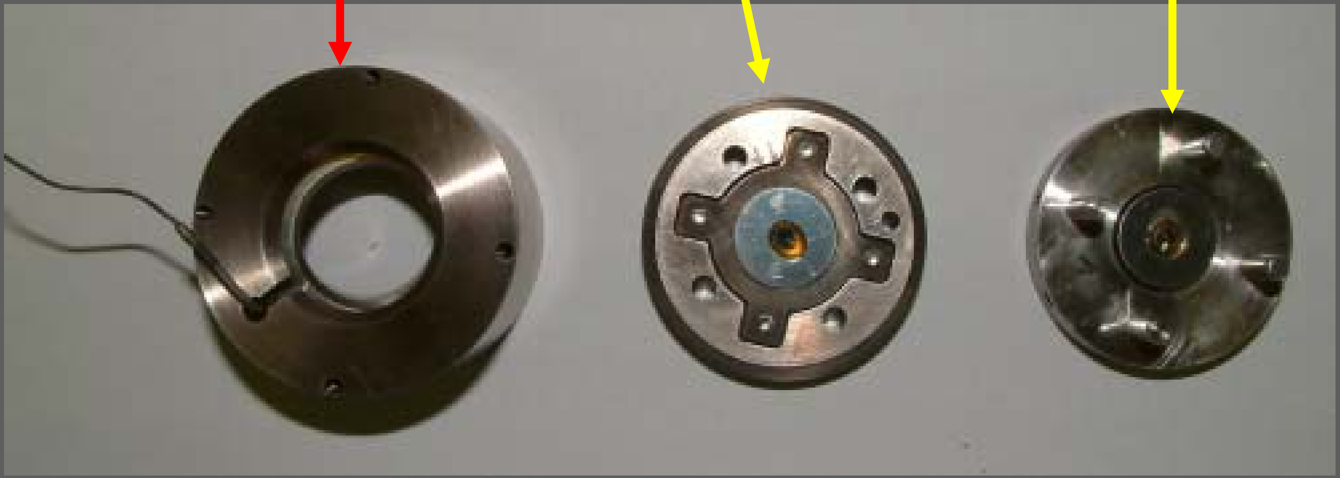
Euphotic zone 0 – 200 m (20 bar)

Bathyal zone 200 – 2000 m (200 bar)

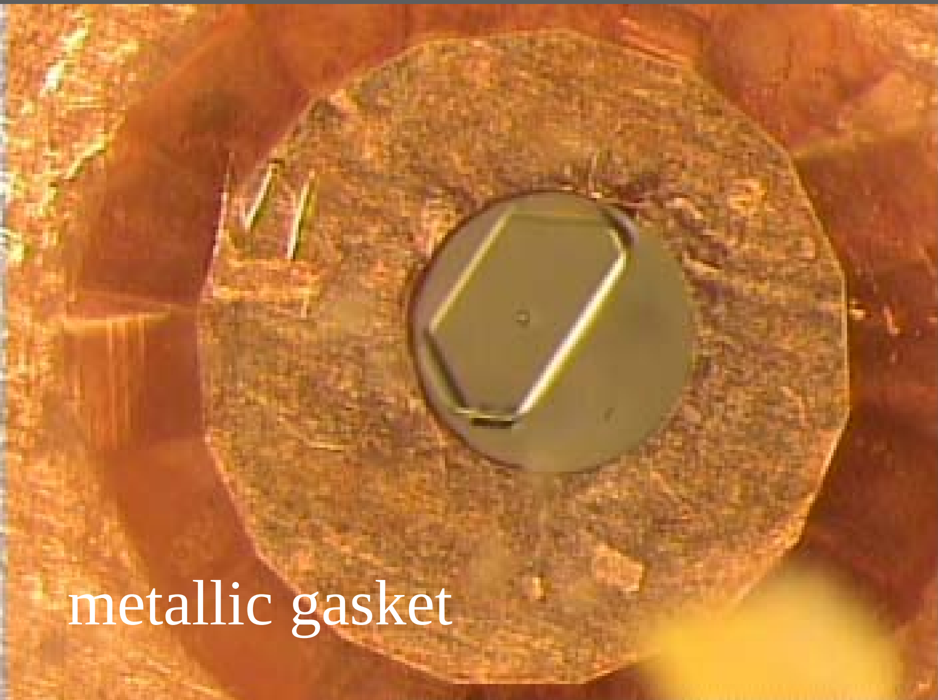
Abyssal zone 2000 – 6000 m (600 bar)

Hadal zone 6000 –11000 m (1.1 kbar)





Diamond anvil cell

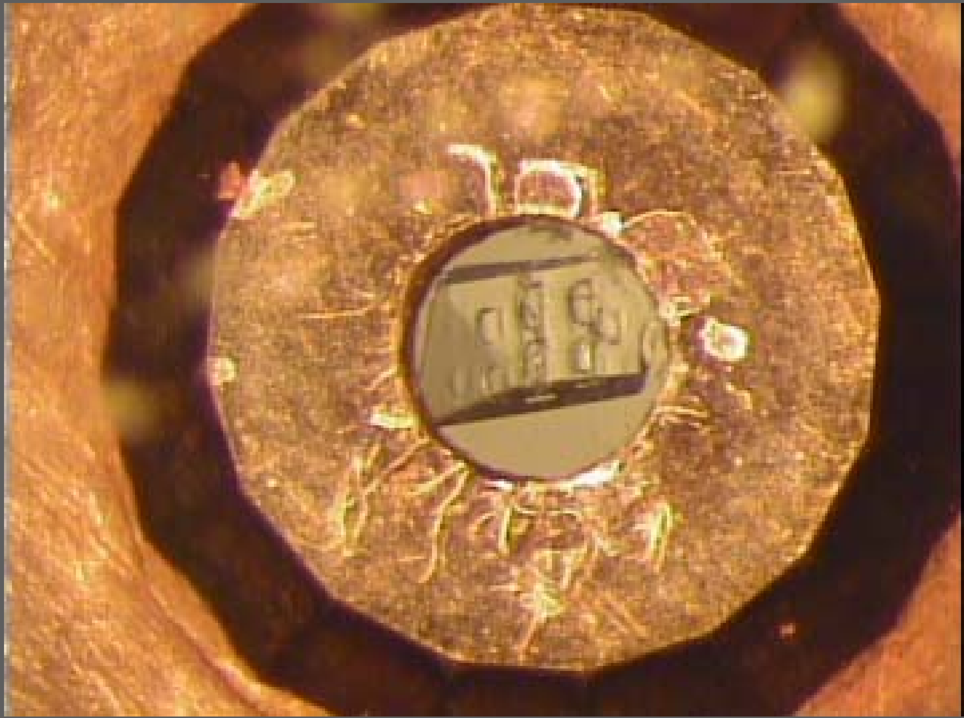


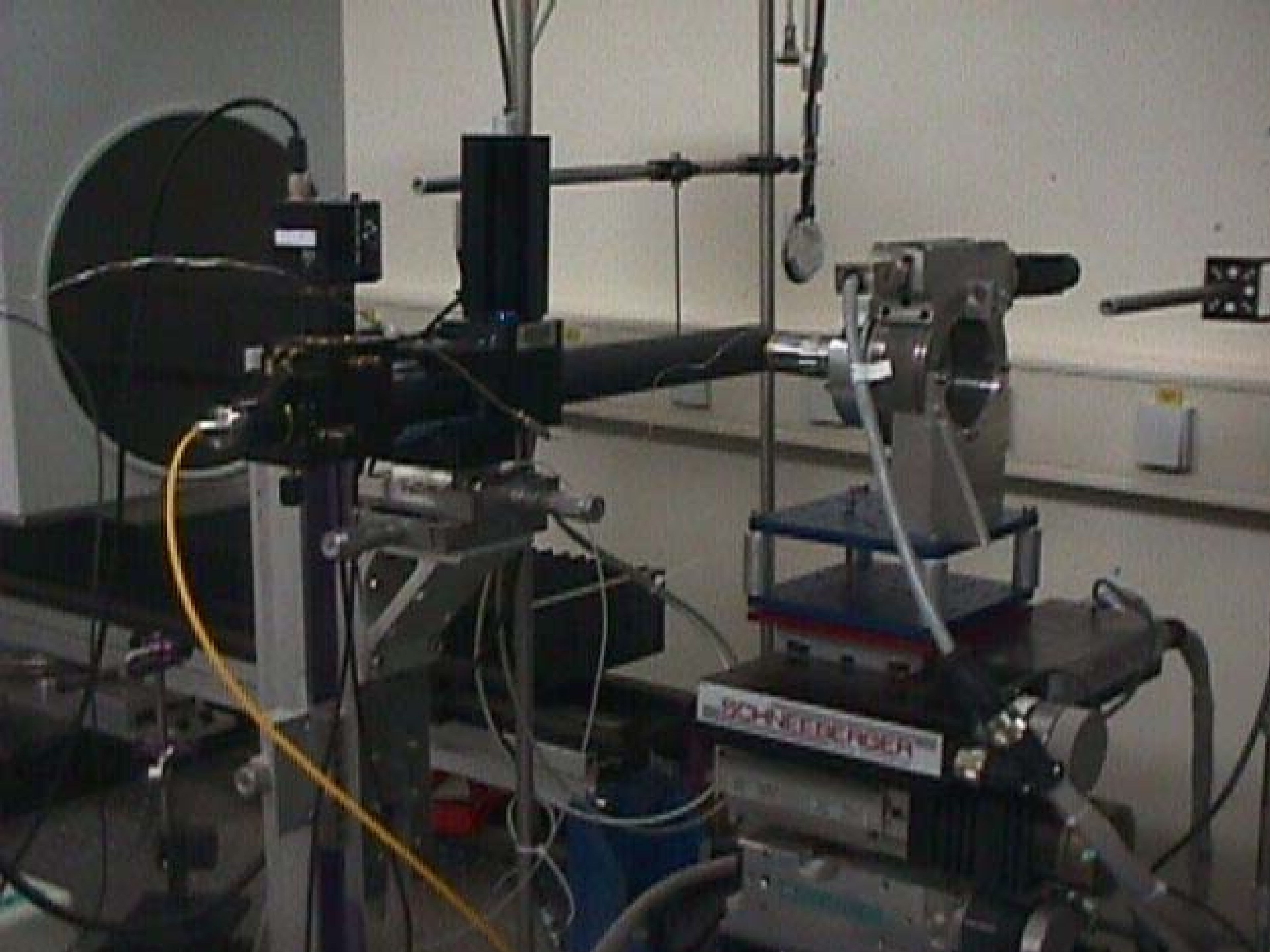
- cavity diameter: 400 μm
- thickness: 150-200 μm
- wavelength: λ = 0.331 Å
- crystal-to-detector: 600-1600 mm
- exposure time: 30 - 90 s/deg.
- room temperature

Fresh urate oxidase crystal

Crystal after multiple irradiations

beam size: 50 x 50 μm^2




SCHNEEBERGER

Advantages of ultra-short wavelengths

- Reduced absorption and extinction by diamonds.

  Absorption of the direct beam by the first diamond is nearly constant for a narrow range of oscillation, and is taken into account by the scaling procedure. Absorption of diffracted beams by the second diamond could be corrected by calculation of the absorption path for each reflection; at the present stage, no correction was applied as the maximum variation of transmission due to path length variations was small at $\lambda < 0.4$ Å.

- Diffraction is confined within a narrow forward cone

  At $\lambda = 0.33$ Å, the full opening angle $4\theta_{max}$ is 25.3° at 1.5 Å resolution; thus the limited useful aperture of DAC does not limit resolution.

- Long crystal-to-detector distance

  Diffracted X-rays are nearly normal to detector

- Less radiation damage?

X-ray beam characteristics

- **Parallel beam**

  required for small Bragg-spot size
- **Bright source**

  required since scattered intensity varies as λ^2

Unfocused radiation from in-vacuum narrow-gap undulator

- **Choice of wavelength**

  Imaging plate detector has a sensitive layer composed of BaFBr: Eu^{2+}.

  The detective quantum efficiency (DQE) of the detector is governed by the stopping power of Ba.

Near the Ba K-absorption edge ($\lambda \leq 0.3311$ Å)

Higher signal

DQE is maximal for elastic scattering.

Lower noise

Noise is mainly produced by **Compton scattering** from diamonds.

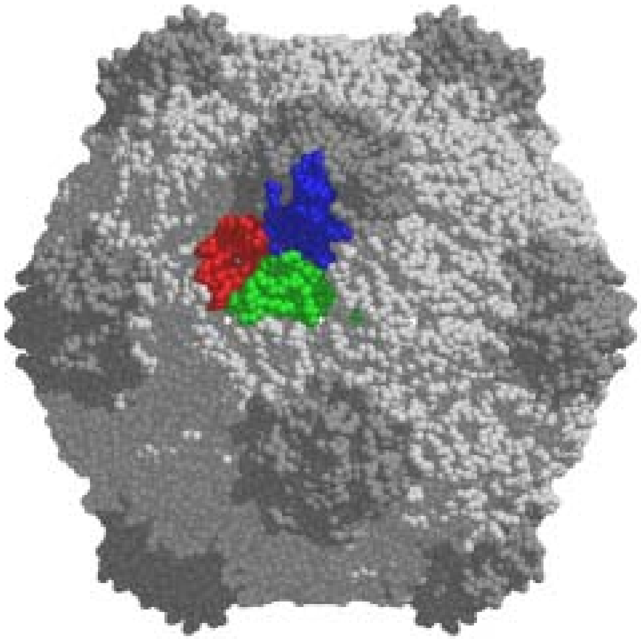
The **energy shift** of Compton scattering is given by $\Delta E = -h^2/(2md^2)$ where h is the Planck's constant, m is the rest-mass of electron and $d = \lambda/(2\sin\theta)$ is the crystallographic resolution.

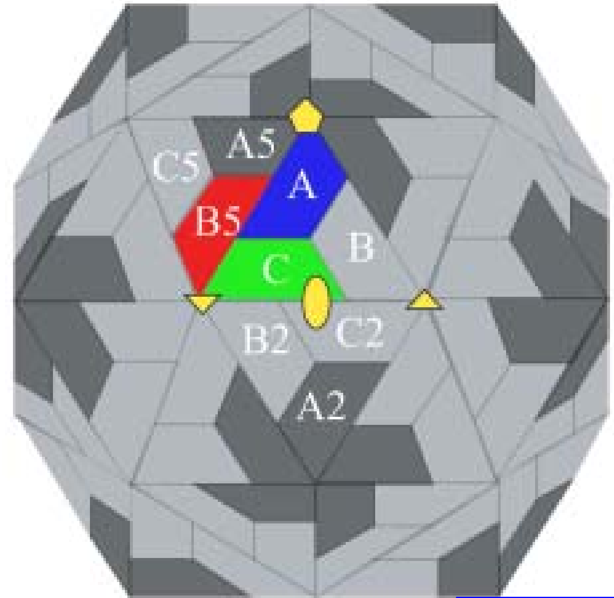
Although small at low resolution (6 eV at 5Å resolution), this shift is sufficient at medium (16.7 eV at 3 Å resolution) and high resolution (66.8 eV at 1.5 Å resolution) to produce a significant variation of the barium absorption coefficient. This effect thus **improves the signal-to-background ratio where most useful**, i.e. in the region of the diffraction pattern where signals are the weakest.

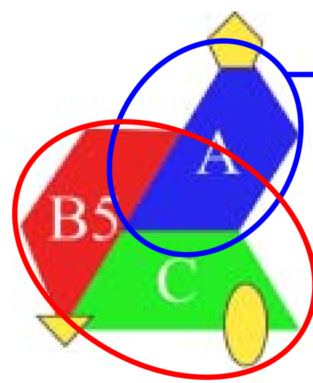
Cowpea mosaic virus

- Cowpea mosaic virus (CPMV): a member of the **Comovirus** family, a group of **icosahedral plant viruses** similar to mammalian *Picornaviruses*. Comoviruses have a bipartite, single stranded RNA genome and each RNA molecule is encapsulated in a separate particle.

- Crystals with rhombic dodecahedral morphology in space group I23 (a = 317 Å) used for the first structure determination at 3.5Å resolution. Refinement at 2.8 Å by Lin et al.







Small protein (S):
189 amino acids
1475 non-H atoms
A domain

Large protein (L):
369 amino acids
2865 non-H atoms
B+C domains

Polymorphism of cubic CPMV crystals at atmospheric pressure

• Only about 10 % of the rhombic dodecahedral CPMV crystals, grown in standard conditions, conform strictly to the characteristics of cubic **I23** space group. The other crystals have weak reflections with indices of h + k + l = 2n + 1 (odd reflections), which should be absent in the I23 space group. Space group **P23** was tentatively assigned to these crystals, which agreed with the presence of odd reflections but **not with the packing** criteria of the virus particles. The intensities of odd reflections varied from one crystal to another and their intensity increased from low to high resolution. The crystals with odd reflections are **poorly organized** and diffract X-rays to low resolution.

Space group: I23
Diffraction: 2.8 Å

10 % ←

→ **90 %**

Space group: P23
Diffraction: 4 Å

• The relationship between the I23 and P23 cells is interpreted as a **first order phase transition**. With respect to the standard positions in the I23 structure, virus particles are rotated randomly with an angle of approximately 7.6°. This model is compatible with the abnormal behaviour of intensities with resolution.

Effect of pressure on cubic P23 CPMV crystals

Rhombic dodecahedral crystals mounted in stabilizing mother liquor

Aqueous solution of MPD, ammonium sulfate, PEG 8000, potassium phosphate buffer pH 7

Data collection parameters

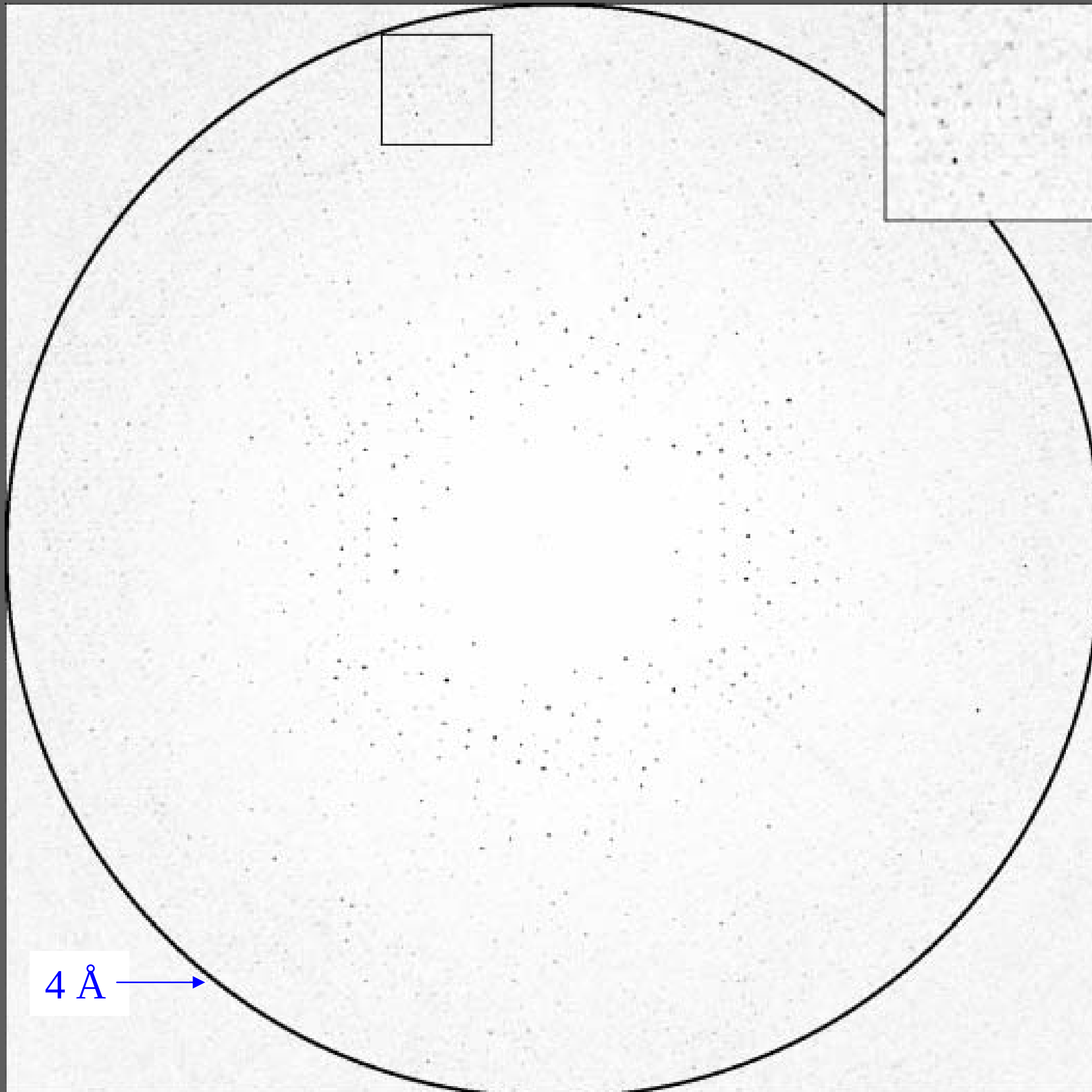
Wavelength: λ = 0.331 Å

Beam cross-section: 50 µm x 50 µm

Oscillation range: 0.3°

- At 2.0 kbar: **cubic P23**, a = 315.1 Å, mosaicity: ~ 0.3°, **resolution: ~ 4-5 Å**
- At 3.3 kbar: **cubic I23**, a = 314.0 Å, mosaicity: ~ 0.08°, **resolution: ~ 2.7 Å** reflections visible to ~ 2.2 Å.
- At 4.4 kbar: **crystal destroyed**.

R. Fourme, I. Ascone, R. Kahn, M. Mezouar, P. Bouvier, E. Girard, T. Lin & J. E. Johnson (2002), Structure ***10****, 1409.*

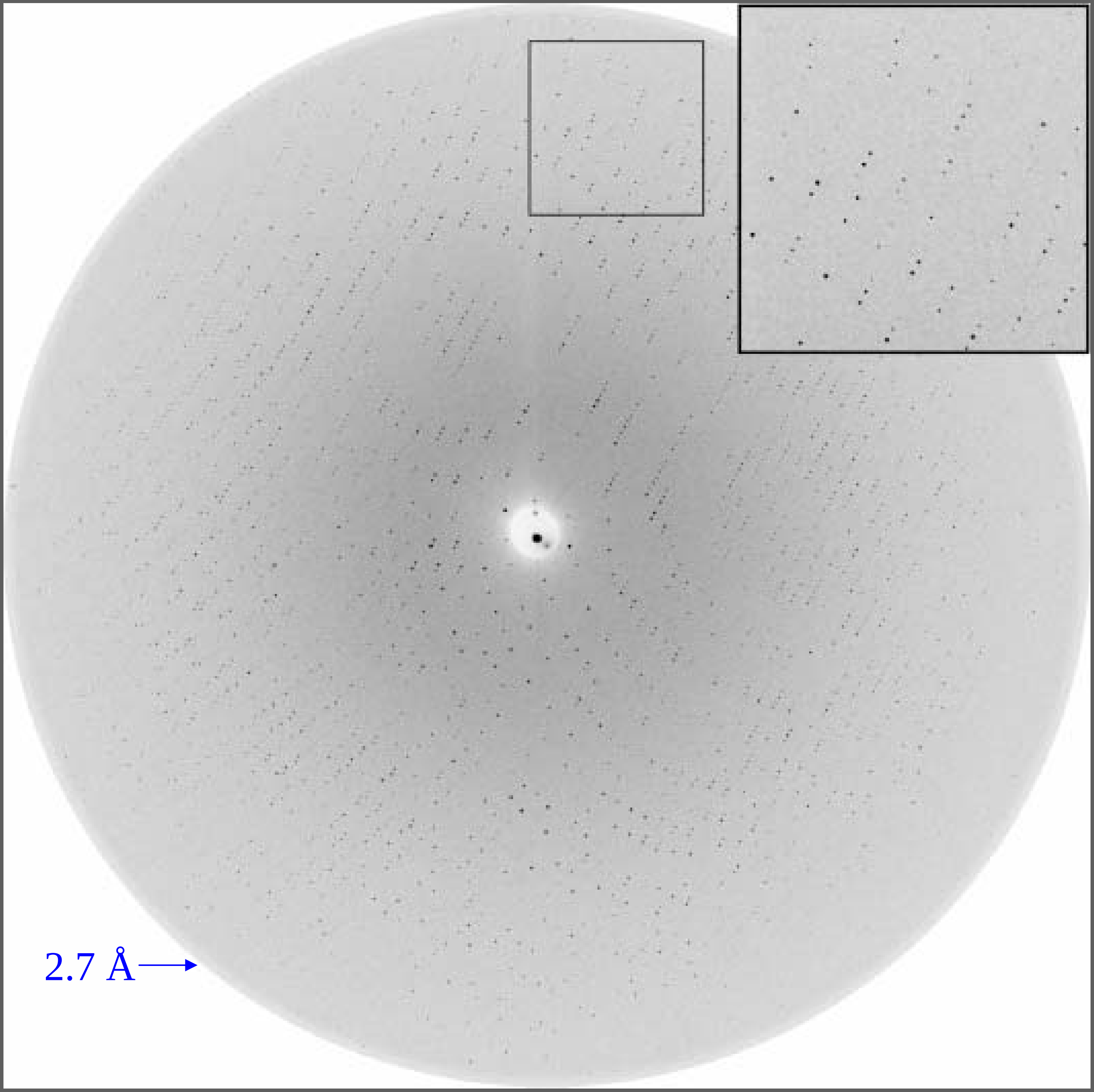


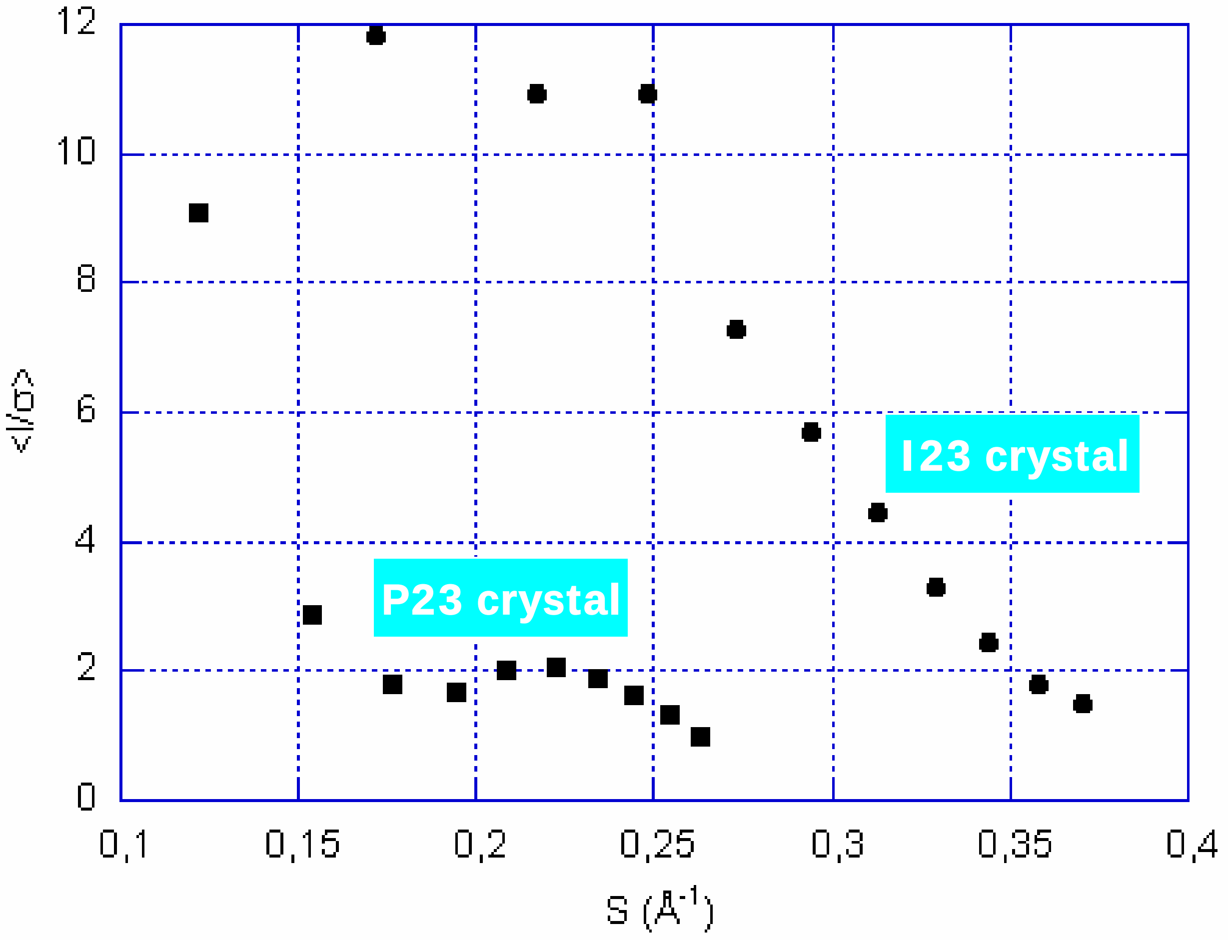
P23
2 kbar
osc. 0.3°

After transition

I23
3.3 kbar

osc. 0.3°
exp. 90 s.





CPMV: average s/n of diffraction
at 2.0 kbar (squares)
at 3.3 kbar (circles)

CPMV: data collection at 3.3 kbar

- Nov. 30-Dec. 2/ 2002 ; station ID30 at the ESRF ; Si(111) monochromator ; λ = 0.331 Å ; unfocused beam 50 x 50 μm^2 ; MAR345 imaging plate ; crystal-to-detector distance: 1375 mm ; ESRF: 2/3 filling mode, <I> ~ 180 mA.
- Exposure time: 90 s/0.3 deg. (5 min./deg.)

  will be reduced to about 40 s/deg. after transfer to ID27 with two in-vacuum U23 undulators
- Typically 1.5-2 deg./irradiated zone ;

  4 irradiated zones /crystal → **6-8 useful deg./sample at room temperature**
- 8 crystals
- Data completeness: 91%
- Resolution: 2.8 Å (I/σ cutoff = 2.)

The I23 structure at atmospheric pressure that is used for comparison was refined at 2.8 Å by Johnson et al. (data collected using A1 station at CHESS, Cornell University)

Dmin (Å)	Rfull %	Rcum %	< I >	σ	I/σ	Nmeas	Nref	Compl. %	Mult.
8.84	5.7	6.1	1883	166	11.3	11328	3021	89.0	3.3
6.25	10.4	8.9	1130	174	6.5	21950	5586	91.2	3.5
5.10	10.8	10.2	1301	209	6.2	28712	7251	91.0	3.5
4.42	8.1	9.5	2184	261	8.4	33668	8594	91.4	3.5
3.95	8.9	9.6	2097	277	7.6	38248	9710	91.6	3.5
3.61	12.0	10.3	1525	271	5.6	42471	10715	92.0	3.5
3.34	15.4	11.3	1223	274	4.5	46395	11747	92.2	3.5
3.13	20.4	12.5	904	271	3.3	49546	12536	91.8	3.5
2.95	25.9	13.8	700	263	2.7	51586	13208	92.0	3.4
2.80	34.4	14.9	487	242	2.0	46696	12837	88.6	3.1
Overall	13.8	14.9	1248	254	4.9	370600	95205	91.2	3.4

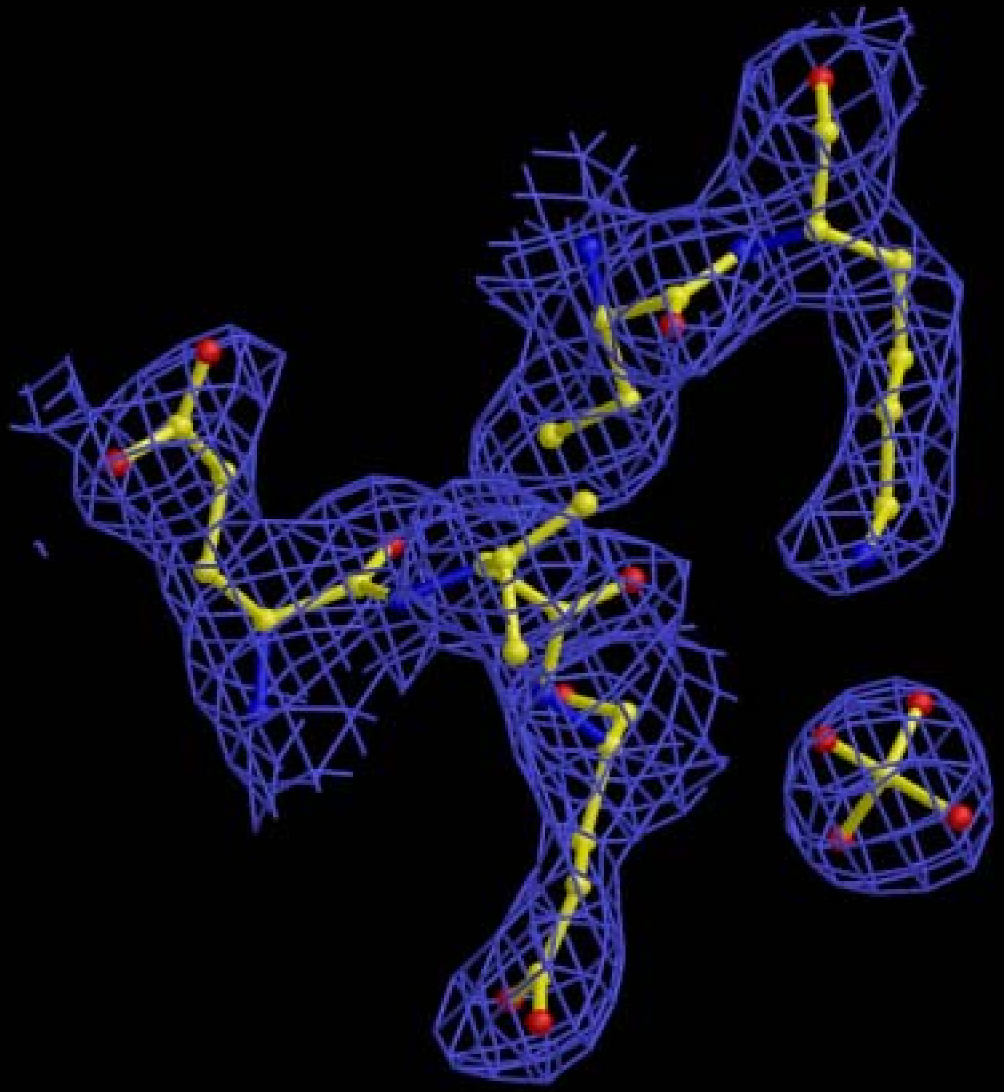
CPMV at 3.3 kbar
Summary of data reduction

Rfull: Rsym in the shell for fully recorded reflections ; Rcum: cumulated Rsym for all reflections ;
Nmeas: number of measurements ; Nref: number of unique reflections ; Compl.: completeness ; Mult: multiplicity.

Measurements: 370600 (91.2 %)
Rejected reflections: 3277 (1.0 %)

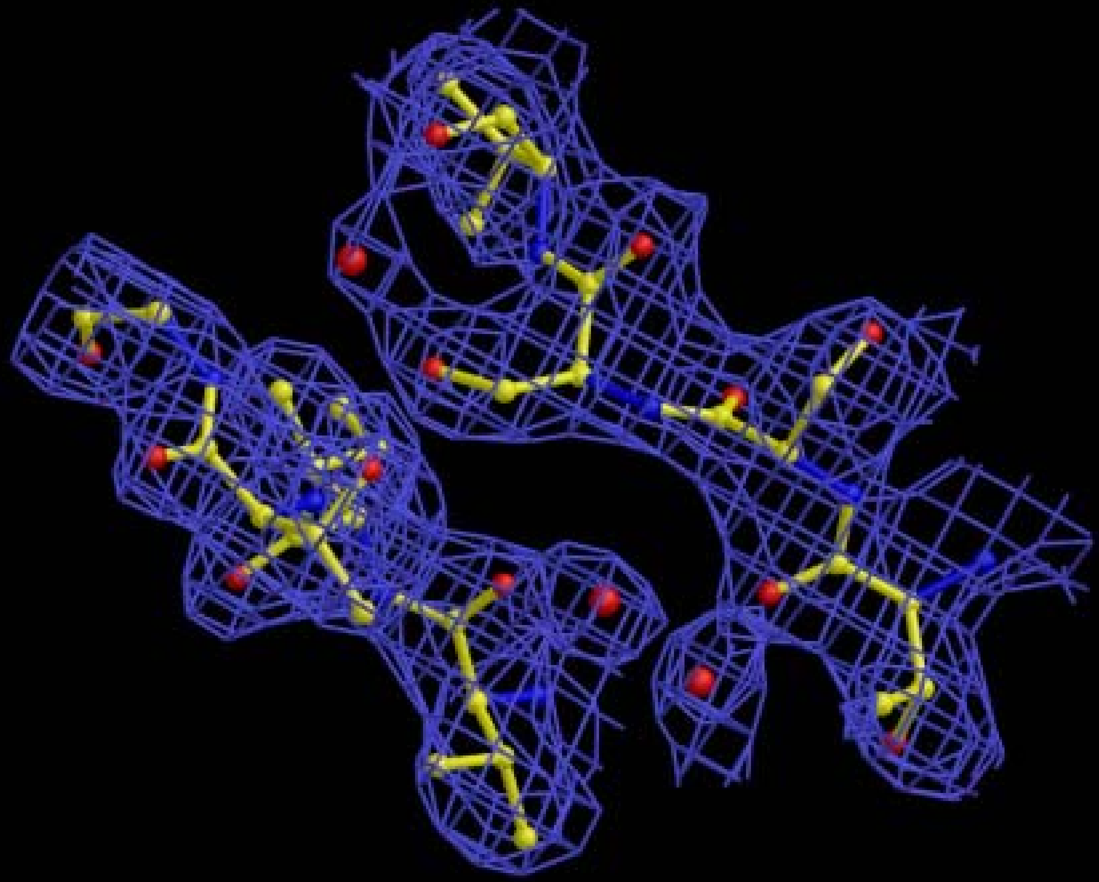
Fully recorded reflections: 238288

Measurements of partials: 132312



Interface between the two domains of protein L

(Cys+Lys / Glu+Val+Glu)



Zone at the interface between domain C of protein L and domain B of a neighboring L protein

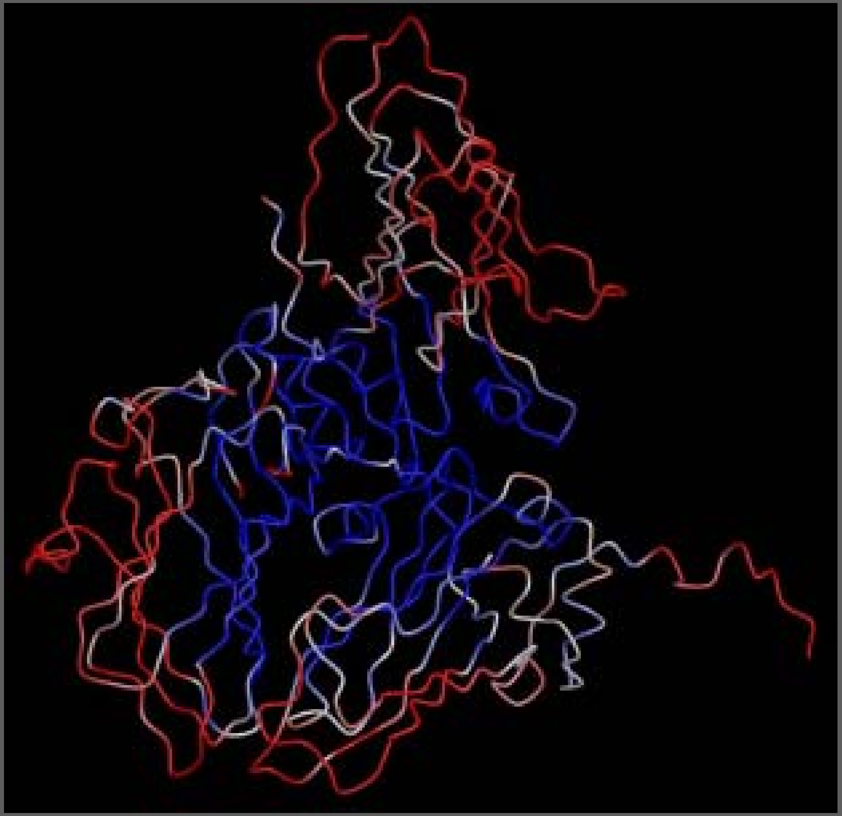
Three ordered water molecules (red) are visible.

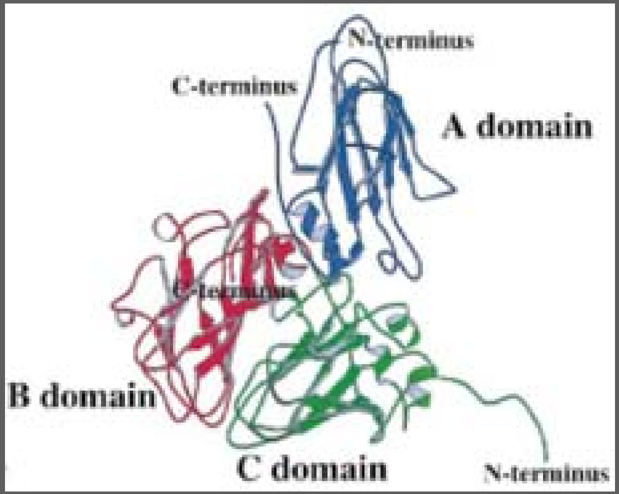
- **Reduction of cell volume: 3.4 %**

    Structure at atmospheric pressure: a = 317.0 Å

    Structure at high pressure: a = 313.4 Å

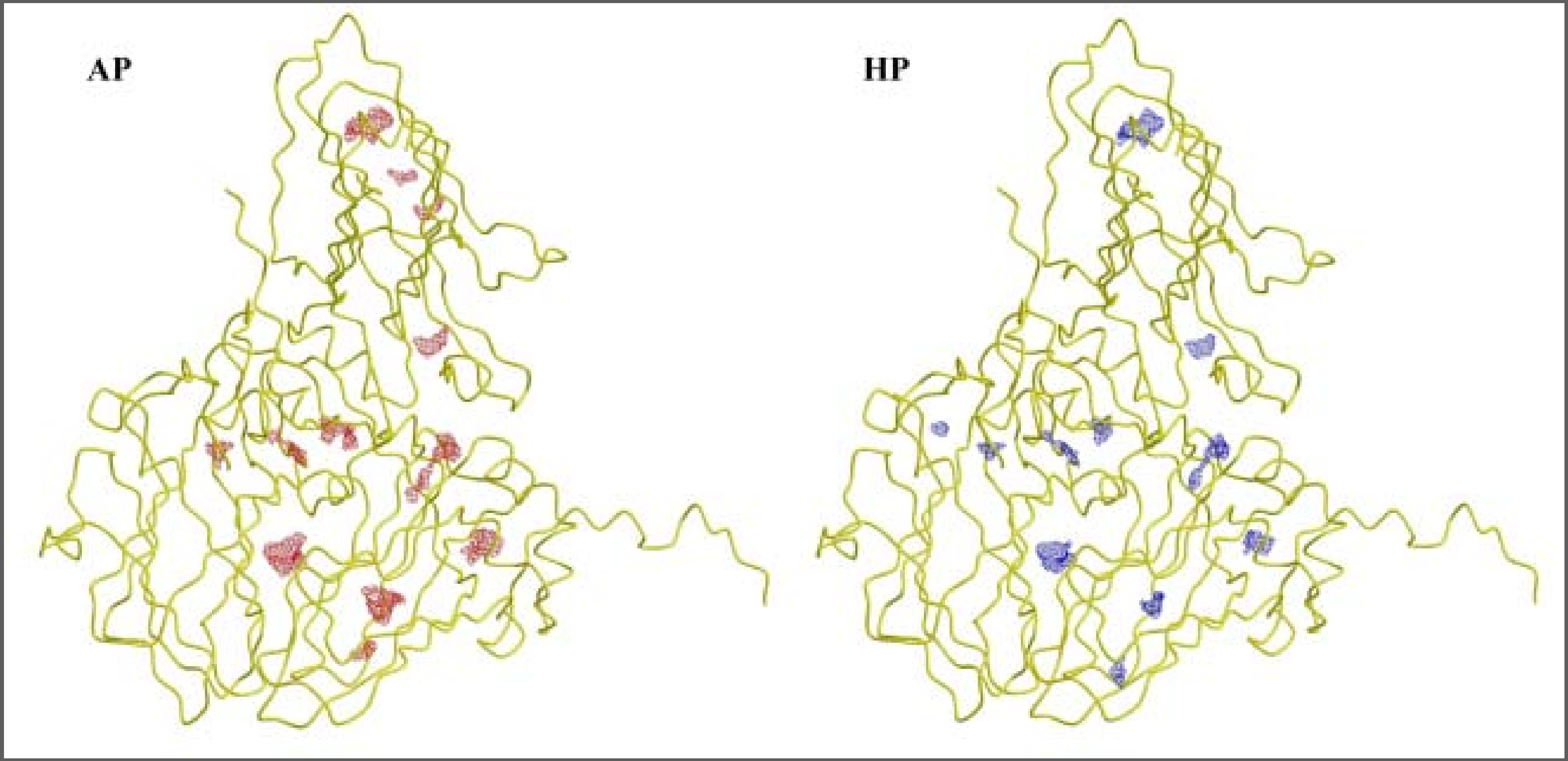
- **Reduction of capsid volume: 3.7 %**
- **Structure superposition**





CPMV: evolution of cavities

- Reduction in the number of cavities
- Reduction of cavity volumes by about 40 %



CPMV: evolution of hydrogen bond lengths

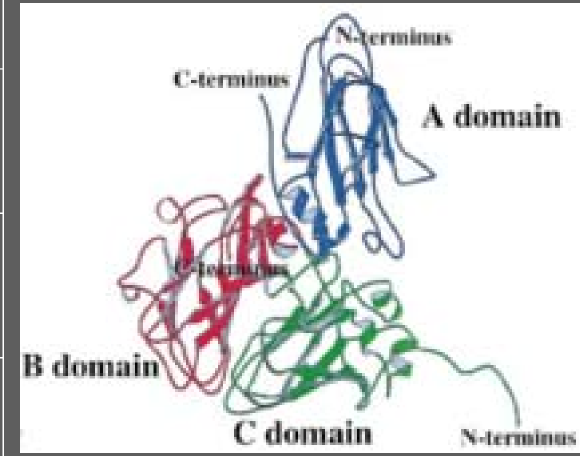
Hydrogen bond nature	No. of measurements	HP structure		AP structure	
		Mean value (Å)	Standard deviation	Mean value (Å)	Standard deviation
Protein - protein	428	2.92	0.19	2.95	0.19
Protein - protein in S-domain	145	2.93	0.23	2.96	0.17
Protein - protein in L-domain	268	2.91	0.20	2.95	0.20
Protein - water molecule	171	2.87	0.21	2.92	0.22
Water - water molecules	24	2.84	0.17	2.94	0.25

CPMV: evolution of ordered water molecules

- Atmospheric pressure structure: 99
- High pressure structure: 195

CPMV: evolution of temperature factors

Mean B factor (Å^2)		HP structure	AP structure
Small and Large proteins		20.30	29.17
	Main chain	19.75	28.78
	Side chains	20.84	29.54
A domain		24.01	32.56
	Main chain	23.73	32.44
	Side chains	24.21	32.61
B5 domain		19.63	28.05
	Main chain	19.29	27.85
	Side chains	19.98	28.20
C domain		17.08	26.77
	Main chain	16.08	25.94
	Side chains	18.16	27.68
Water molecules		24.84	30.26

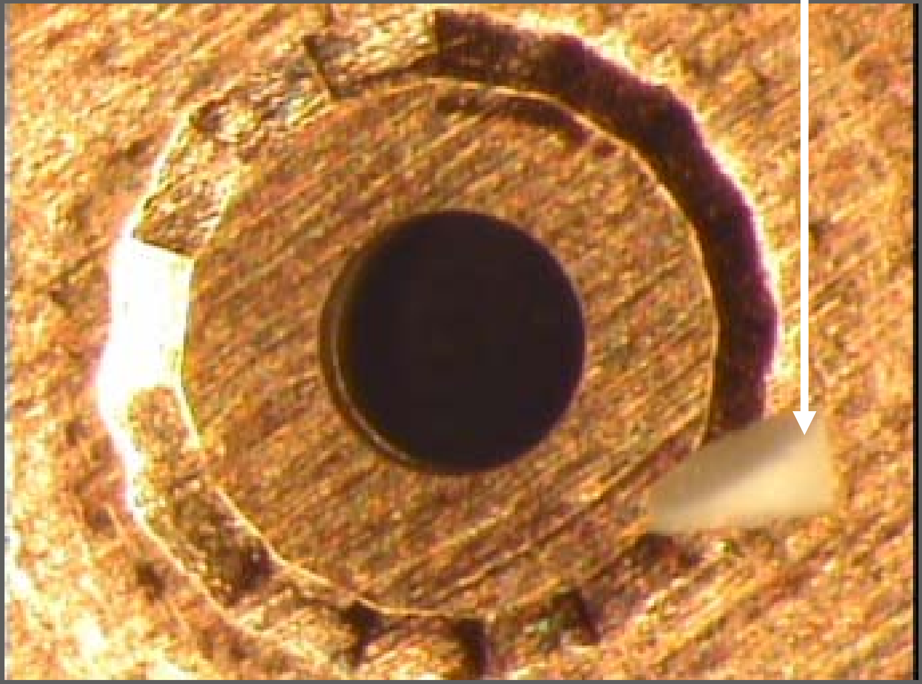


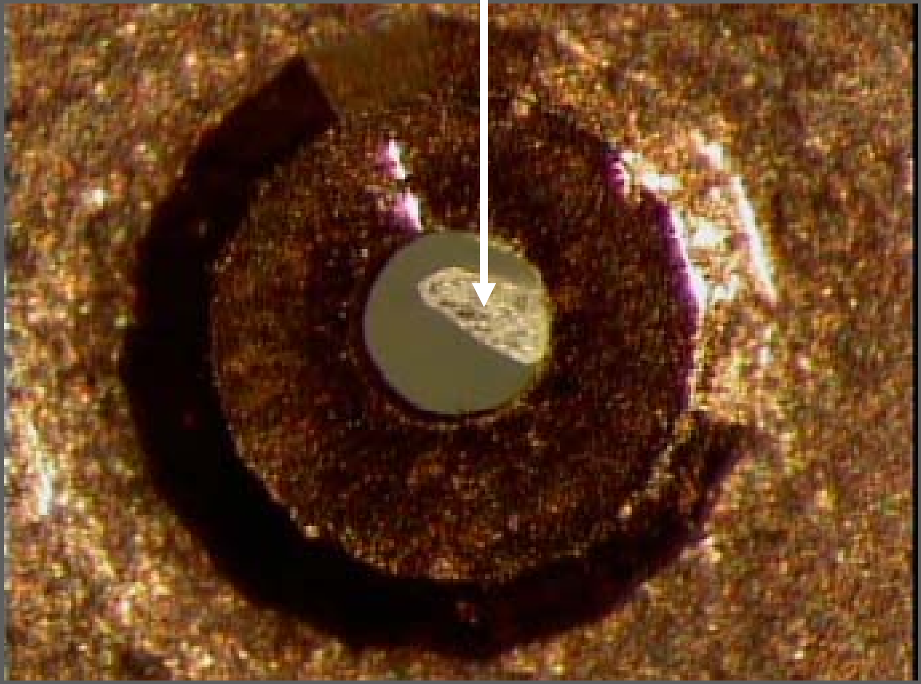
Recent improvements on diamond anvil cells

- Larger opening (42° → 62° → 90°)
- Piston with reduced friction → improved reversibility
- Softer gaskets (Cu, Pd instead of Inconel) → improved pressure stability
- Low Z (BN, C) splinters forcing crystal to reorient

Micromachined BN piece

Diamond splinter





Urate oxidase crystals

Space group I222
High completeness data collection from three crystals using diamond splinter

D_{min} (Å)	R_{mrg}	I/sigma	Nmeas	Nref	Completeness (%)	Multiplicity
7.25	0.037	14.2	1809	516	90.8	3.3
5.12	0.040	15.6	4147	975	96.8	4.0
4.18	0.033	19.2	5301	1239	98.2	4.0
3.62	0.040	15.2	6487	1485	98.4	4.1
3.24	0.052	12.4	7385	1665	97.6	4.2
2.96	0.075	9.1	8447	1870	99.4	4.3
2.74	0.108	6.3	9182	2018	98.4	4.3
2.56	0.160	4.2	9937	2169	98.6	4.4
2.42	0.209	3.1	10673	2332	99.9	4.4
2.29	0.255	2.6	10575	2322	95.0	4.3
Overall:	0.068	9.2	73943	16591	97.8	4.2

Fully fledged HP macromolecular crystallography

Potential of HP crystallography

- Is the **ordering effect** of high pressure more **general** than just the particular CPMV case?
  if yes, pressure as an alternative to space-grown crystals ?

- Pressure as a way to modify the **Gibbs free energy**:
  exploration of phase transitions and protein substates.

- Structure of proteins from **extremophiles**

- Detailed analysis of **interactions** :
  - between protein molecules within the crystal
  - between monomers of an oligomeric protein
  - between components of a complex macromolecular assembly
  - water-water and water-molecules interactions

- First steps of **protein denaturation** induced locally by pressure.

- Determination and refinement of parameters based on structural information (solvent accessible surface, packing…) for the **estimation of thermodynamic quantities** such as free enthalpy.

Data collection efficiency (DCE)

• The ratio of diffracting volumes required to acquire a high completeness data set at a given resolution might be used to compare the efficiency of data collection in different experimental conditions. Instead, we suggest to use a definition of data collection efficiency (DCE) based on subsets of data and thus unbiased by the particular course and hazards of a full data collection. DCE is defined as $d\omega/dv$, where $d\omega$ is the angular range throughout which useful data can be acquired from an irradiated zone of the crystal and dv is the volume of this zone. DCE is the relevant parameter for the crystallographer, as it quantifies the intuitive notion of "taking the best from a given sample". It is shown that the DCE is larger for the HP case.

• Note that in our experiments on CPMV crystals, many parameters are involved in the improvement of DCE and it is difficult to assess the particular role of one of them. For instance, we cannot conclude from our results that ultra-short wavelengths increase the sample lifetime.

Parameters for CPMV data collection

AP set: 279 K and atmospheric pressure at CHESS
HP set: 296 K and 3.3 kbar at ESRF

Common specifications and parameters

	AP	HP
Beamline	CHESS A1	ESRF ID30
Wavelength	0.908 Å	0.3305 Å
Detector	Fuji IP (BAS 2000)	Fuji IP (Mar 345)
Beam cross section (µm)	100 x 100	50 x 50
Oscillation angle per image (deg.)	0.3	0.3
Exposure time per image (s)	45	90

Parameters for the selected crystal

	AP	HP
Crystal size (µm)	~1000 x 1000 x 1000	~170 x 170x 170
Initial mosaicity (deg.)	0.080	0.045
Number of irradiated zones	9	3
Volume of irradiated zone (µm)	~100 x 100 x 1000	~50 x 50 x 170
Total useful angular range (deg.)	18	7.2
Largest angular range acquired per irradiated zone (deg.)	2.4	3.3
DCE for this zone (deg./mm^3)	240	7700
Relative DCE	1	30
DCE for the whole crystal (deg./mm^3)	18	1450
Relative DCE	1	80

Implications for standard data collection

• The high efficiency of the HP data collection is based on the multi-faceted improvement in signal to noise provided by the particular experimental conditions used in these experiments. We suggest that, on this basis, there is ground for improvement in conventional data collection (i.e. at atmospheric pressure) which would be especially beneficial for difficult problems including small, radiation sensitive and poorly diffracting samples. Another benefit might be to reduce systematic errors and increase the internal consistency of data, which is an issue in particular for the exploitation of small anomalous signals for phasing at a single wavelength.